

Usando Kinect como Sensor para Pulverización Inteligente

C. Correa¹, C. Valero¹, P. Barreiro¹, J. Ortiz-Cañavate¹ y J. Gil¹

¹ Universidad Politécnica de Madrid, Dpto. de Ingeniería Rural. LPFTAGRALIA. Av. Complutense s/n. Ciudad Universitaria, Madrid, ccorrea@udec.cl.

Resumen

Este trabajo está orientado a resolver el problema de la caracterización de la copa de árboles frutales para la aplicación localizada de fitosanitarios.

Esta propuesta utiliza un mapa de profundidad (*Depth image*) y una imagen RGB combinadas (RGB-D), proporcionados por el sensor *Kinect* de Microsoft, para aplicar pesticidas de forma localizada. A través del mapa de profundidad se puede estimar la densidad de la copa y a partir de esta información determinar qué boquillas se deben abrir en cada momento.

Se desarrollaron algoritmos implementados en *Matlab* que permiten además de la adquisición de las imágenes RGB-D, aplicar plaguicidas sólo a hojas y/o frutos según se desee. Estos algoritmos fueron implementados en un software que se comunica con el entorno de desarrollo "*Kinect Windows SDK*", encargado de extraer las imágenes desde el sensor *Kinect*.

Por otra parte, para identificar hojas, se implementaron algoritmos de clasificación e identificación. Los algoritmos de clasificación utilizados fueron "*Fuzzy C-Means* con *Gustafson Kessel*" (FCM-GK) y "*K-Means*". Los centroides o prototipos de cada clase generados por FCM-GK fueron usados como semilla para *K-Means*, para acelerar la convergencia del algoritmo y mantener la coherencia temporal en los grupos generados por *K-Means*. Los algoritmos de clasificación fueron aplicados sobre las imágenes transformadas al espacio de color $L^*a^*b^*$; específicamente se emplearon los canales a^* , b^* (canales cromáticos) con el fin de reducir el efecto de la luz sobre los colores. Los algoritmos de clasificación fueron configurados para buscar cuatro grupos: hojas, porosidad, frutas y tronco. Una vez que el clasificador genera los prototipos de los grupos, un clasificador denominado Máquina de Soporte Vectorial, que utiliza como núcleo una función Gaussiana base radial, identifica la clase de interés (hojas).

La combinación de estos algoritmos ha mostrado bajos errores de clasificación, rendimiento del 4% de error en la identificación de hojas. Además, estos algoritmos de procesamiento de hasta 8.4 imágenes por segundo, lo que permite su aplicación en tiempo real.

Los resultados demuestran la viabilidad de utilizar el sensor "*Kinect*" para determinar dónde y cuándo aplicar pesticidas. Por otra parte, también muestran que existen limitaciones en su uso, impuesta por las condiciones de luz. En otras palabras, es posible usar "*Kinect*" en exteriores, pero durante días nublados, temprano en la mañana o en la noche con iluminación artificial, o añadiendo un parasol en condiciones de luz intensa.

Palabras clave: RGB-D, Kinect, pulverización inteligente.

Using Kinect as sensor for smart spraying

Abstract

This work is aimed to solve the problem of fruit tree canopy characterization for spraying purposes. This proposal is based on using a depth map and a RGB image provided by the vision sensor "*Kinect*" from Microsoft to perform smart spraying.

Through the depth map the density of trees can be estimated and thereby it is possible to determine which nozzles should be turned on/off at every moment. Furthermore, algorithms to apply pesticides only to leaves and/or fruits as desired were created.

The software was developed in *Matlab* that allows the acquisition of the depth map and RGB image from the "*Kinect*" sensor. This software communicates with the "*Kinect Windows SDK*", processes the information, and then provides information regarding the presence and location of leaves and/or fruit.

To identify leaves, classification and identification algorithms were applied. The classification algorithms used were "Fuzzy C-Means with Gustafson Kessel" FCM-GK and "K-Means". In this framework the centroids generated by FCM are used as seed for K-means, in order to accelerate the implementation and maintain temporal consistency in the groups generated by the K-means algorithm. Classification algorithms were applied over the images transformed to the $L^*a^*b^*$ color spaces, specifically over the channels a^*b^* (the chromatic channels) in order to reduce the light effect over the colors. Classification algorithms were tuned to search for four clusters: leaves, porosity, fruits, and trunk. Once the classifier generates the cluster prototype, leaves are identified by using a binary Support Vector Machine that uses as kernel a Gaussian radial basis function.

The combinations of all these algorithms have shown low misclassification, yield 4% error on the leaves identification. Besides, these algorithms process up to 8.4 frames per second, allowing its application in real time.

Result shows the feasibility of using the "Kinect" sensor to determine where and when to apply pesticide. On the other hand, it also shows the limitation imposed by the lightening conditions. In other words, it is possible to use "Kinect" outdoors, but working during cloudy days, early in the morning or at night using artificial illumination, or adding a sun shield for strong light conditions.

Keywords: RGB-D, Smart Spraying, Kinect.

Introducción

La aplicación de fitosanitarios es una práctica usual en la agricultura. Según la FAO "un factor decisivo de la Revolución Verde ha sido el desarrollo y aplicación de plaguicidas para combatir una gran variedad de plagas insectívoras y herbáceas que, de lo contrario, disminuirían el volumen y calidad de la producción alimentaria" (Ongley, 1996). Pero esta práctica cultural no está exenta de problemas de carácter medio ambiental. Esta actividad arroja compuestos orgánicos volátiles a la atmosfera, a la vez que, tanto sus compuestos activos como sus aditivos, contaminan las aguas subterráneas. De ahí el esfuerzo sistemático de los últimos años por intentar reducir su uso y sus efectos cuando son aplicados.

En este sentido es que nacen técnicas, implementos y equipos orientados reducir los volúmenes de aplicación y a mejorar sus condiciones de aplicación (Siegfried et al., 2007). Ejemplo de aquello son los equipos de trazabilidad (Pontikakos et al., 2012; Reyes et al., 2012), los de aplicación selectiva basados en ultrasonido (Brown et al., 2008; Gil et al., 2007), en LIDAR (Walklate et al., 2002) e imágenes (Correa et al., 2012; Diago et al., 2012). Si bien es cierto, los equipos basados en LIDAR tienen un desempeño adecuado, los costos de implementación son elevados (cerca de €3500), lo que relega al LIDAR sólo a aplicaciones en investigación.

Desde esta perspectiva, las aplicaciones basadas en imágenes resultan atractivas, debido a que no son costosas, pero si demandantes de algoritmos capaces de discriminar entre copa y vegetación circundante. En el caso de vegetación distinta a la del cultivo de interés, ésta puede ser diferenciada por los algoritmos (Correa et al., 2012), pero cuando corresponde a parte de la copa de hileras paralelas, los algoritmos no pueden diferenciarlas. Para resolver este problema, en esta investigación se propone el uso de una cámara de uso doméstico RGB-D (de imágenes visibles y de profundidad) que permita discriminar por distancia y color las hileras a pulverizar de las paralelas a la aplicación.

Material y Métodos

La identificación y selección de características en tiempo real de los árboles se realizó mediante el procesamiento de imágenes. Estas imágenes son tanto visibles (RGB), como de profundidad (D). Para la adquisición de dichas imágenes se empleó una cámara RGB-D de *Microsoft Corporation* modelo *Kinect Xbox 360*. Esta cámara, diseñada para su uso en video juegos domésticos, que tiene un costo de €100, provee de video RGB de 1280x1024 pixeles a 30 Hz, así como video infrarrojo (IR) y de profundidad de 640 × 480 pixeles a 30 Hz. Esta cámara se montó en la parte delantera de una unidad móvil automatizada basada en el tractor New Holland modelo Boomer 3050, operado remotamente desarrollado al interior del proyecto europeo RHEA. La cámara se situó a una altura de 1 m sobre el

suelo y perpendicular al cultivo, como se observa en la Figura 1. *Kinect* cuenta con una motorización que le permite, al comienzo de la adquisición de datos, orientarse de forma perpendicular al suelo.



Figura 1. Montaje de la cámara RGB-D (Kinect) en el frontal del tractor.

Las imágenes fueron adquiridas en el campo experimental de Arganda del Rey del CESIC-CAR entre las 16:00 y las 19:00 del 20 de enero de 2013. El cultivo fotografiado fue olivos de 5 m de altura.

Kinect no posee capacidad de almacenamiento de imágenes, por lo que debe ser conectada a un PC que posea el *Kinect Windows SDK* y algún software habilitado para la adquisición. Así con el fin de almacenar y procesar las imágenes se empleó una rutina desarrollada en el entorno de diagramas de bloque, Simulink de Matlab llamada “*Simulink Support for Kinect*” desarrollada por (Chikamasa, 2012). Las imágenes fueron procesadas en un ordenador Intel i7 de 2.4 Ghz y 4Gb de memoria RAM usando Matlab 2012b.

Algoritmos implementados.

Con el propósito de identificar los árboles y sus características de interés, en este caso identificar hojas, es que se programó un algoritmo de clasificación difusa denominado “*Fuzzy C-Means Gustafson-Kessel*” FCM-GK, en paralelo con el algoritmo *K-means*. Estos algoritmos fueron combinados como en (Correa et al., 2012) configurando los algoritmos para que clasifiquen píxeles en cuatro clases distintas. Luego de clasificados los píxeles, son identificados mediante un clasificador binario, denominado Máquina de Soporte Vectorial (SVM), que utiliza como núcleo una función Gaussiana base radial. SVM debido a que es un clasificador binario permite separar copa y no copa.

Estos algoritmos son aplicados a las imágenes RGB como la de la Figura 2 a, y luego filtrados usando como umbral los datos de distancia provenientes de la imagen de profundidad (Figura 2 b).

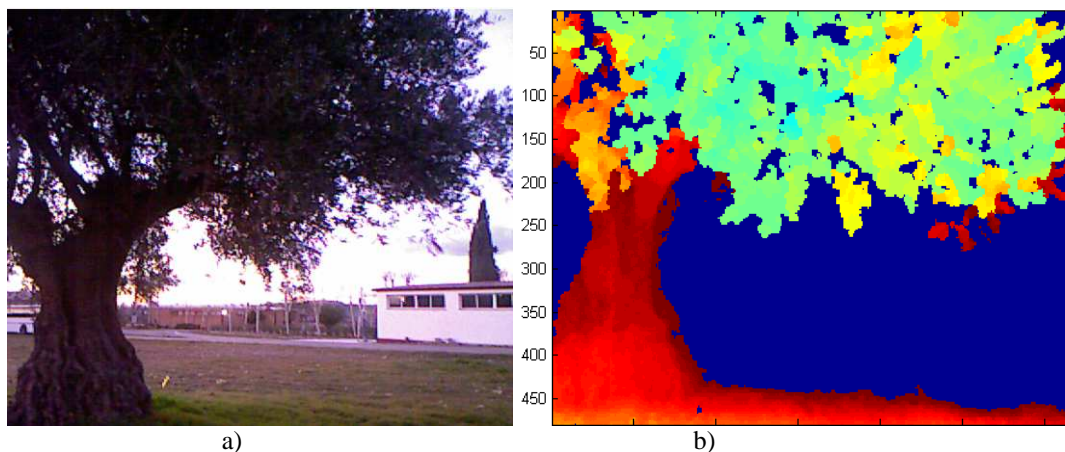


Figura 2. a) Imagen RGB de uno de los olivos. b) Imagen de profundidad generado por Kinect.

En el caso analizado las imágenes RGB y de profundidad son capturadas con una resolución de 480x640. Luego, a partir de la imagen de profundidad, se genera una imagen binaria de 7x13 píxeles, como en la Figura 3 a, que corresponde, en el caso analizado, a los píxeles con distancia menor a 3 m. Si se desea además de la copa pulverizar el tronco esta información puede ser usada directamente para controlar la pulverización (Figura 3 b).

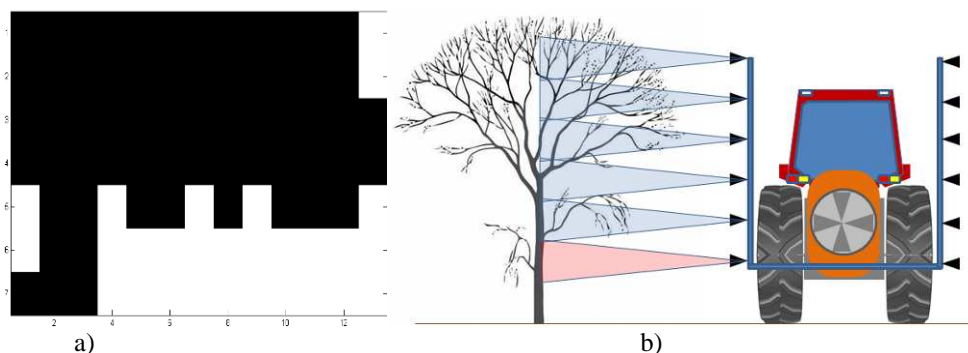


Figura 3. a) Binarización de la presencia de vegetación. b) Esquema del sistema en operación.

En relación al sistema de control de apertura y cierre de válvulas, éste fue simulado usando un microcontrolador Atmega328P a 8 MHz conectado a una batería de relés de 12V y 10A como en la Figura 4. Este microcontrolador recibe, a través de su puerto serial (USART), una cadena de datos que corresponde a las órdenes de encendido y apagado de cada válvula. Estas órdenes son generadas por una rutina creada en Matlab 2012a, que identifica los píxeles de la imagen en base al clasificador SVM descrito anteriormente.

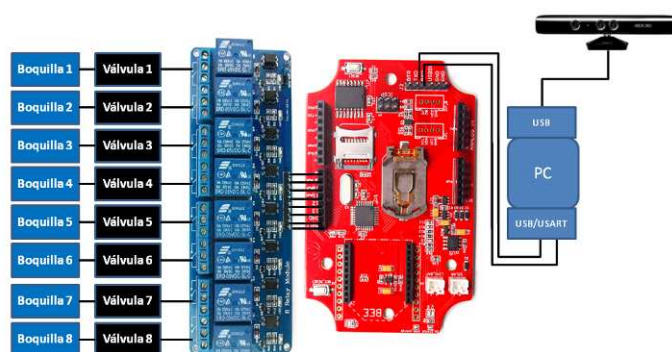


Figura 4. Diagrama de conexión entre Kinect y el sistema de control de válvulas, realizado mediante un ordenador portátil y un microcontrolador Atmega328P a 8 MHz.

Resultados y Discusión

Los resultados pueden ser separados entre aspectos: 1) Error en la detección de la copa. 2) Tiempo de respuesta del sistema. 3) Viabilidad del sistema.

1) Error en la detección de la copa.

Esta a su vez debe ser dividida en detección mediante imágenes RGB, Imágenes de Profundidad y la Interacción de ambas. a) Detección mediante imágenes RGB. En este caso el porcentaje de acierto es del 92% a una resolución de 480x640. Esto, debido a que píxeles perteneciente a árboles distantes en el paisaje, son identificados como píxeles a los que aplicar pesticida. Este problema se ve drásticamente reducido al disminuir la resolución a 80x60 y eliminar zonas no conectadas. Debe notarse que en el caso analizado no existen hileras de árboles paralelas a la hilera procesada. Por lo que los porcentajes de error podrían incrementarse sustancialmente. Al confundir sistemáticamente hojas de la hilera paralela con la hilera de interés.

b) Detección mediante Imágenes de Profundidad. En este caso debido a que la imagen de profundidad tiene un alcance máximo de 4 m, todos aquellos árboles más allá del eje de la hilera procesada, no aparecen en la imagen y por tanto no son procesados. Un ejemplo de aquello puede observarse en la imagen de la Figura 2 b. En este escenario el porcentaje de acierto es del 98%. Si se considera al tronco como objetivo de la aplicación del plaguicida. Si éste no se considera entonces el porcentaje de aciertos cae a un 82%.

c) Detección mediante la interacción de imágenes RGB-D. El problema de la identificación basado únicamente en imágenes visibles presenta la desventaja de no discriminar entre la hilera de interés y las hileras paralelas. Esta desventaja es solventada al usar la información profundidad, ya que sólo son considerados como píxeles válidos en el proceso de clasificación, aquellos que se encuentren dentro de la distancia al eje de la hilera. De este modo el porcentaje de acierto se eleva a un 98%.

2) Tiempo de respuesta del sistema.

El tiempo de respuesta lo podemos dividir en: Tiempo de adquisición de la imagen, Tiempo de procesado de la imagen y Tiempo de actuación de las válvulas. El tiempo de adquisición es de 33 ms (para imágenes adquiridas a 30 fps), el tiempo de procesado es de 100 ms, 80 ms para los algoritmos de clasificación y 2 ms para SVM. Finalmente el tiempo de actuación es de 2 ms, 1 ms en comunicación PC microcontrolador y 1ms en actuación del relé. De este modo el tiempo total desde que se captura una imagen hasta la actuación es de 117 ms, es decir, 8,4 fps. Si consideramos que el campo de visión horizontal de Kinect es de 3 m, el sistema podría actuar hasta velocidades de 25 m s^{-1} .

3) Viabilidad del sistema.

Si bien es cierto que el sistema funciona perfectamente en condiciones de baja iluminación, durante días soleados las imágenes de profundidad no pueden ser generadas debido a que la incidencia directa de los rayos solares sobre las plantas enmascara la matriz de puntos infrarrojos proyectada por la *Kinect* para calcular la profundidad. Por ello se propone incorporar este sistema a pulverizadoras del tipo túnel, ya que su estructura provee sombra artificial, garantizando de este modo la operación a todo evento del sistema. Más importante aún, al conocer la distancia a la copa se puede determinar la dosis a aplicar como en Moltó et al., (2001). Por otra parte, añadiendo una pantalla (Figura 5) que evite “ver” la hilera paralela se puede evitar que los algoritmos la confundan la hilera de interés.

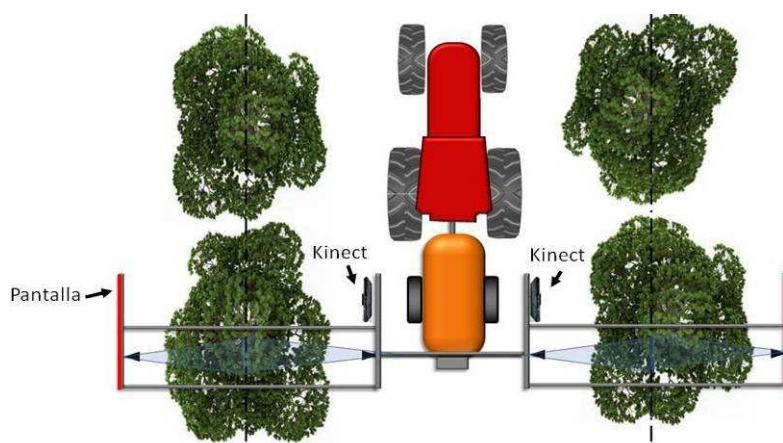


Figura 5. Propuesta de integración en pulverizadoras tipo túnel.

Conclusiones

A pesar de la desventaja que supone que las imágenes de profundidad no estén disponibles en toda circunstancia, debido a que en días soleados la luz IR afecta el desempeño de la cámara, estas ofrecen una importante mejora frente a la alternativa de usar sólo imágenes visibles. Es más, la combinación

RGB-D ha mostrado mejoras significativas en la identificación y localización de la copa de los olivos. Se visualizan dos aplicaciones futuras de este tipo de cámaras; la primera en aplicaciones robotizadas, que pueden operar de noche, y en pulverizadoras tipo túnel en las que el túnel cumple una doble función, evitar la deriva y recuperar el exceso de plaguicida y por otra parte desde la perspectiva de esta aplicación, sirve como pantalla de bloqueo visual de la hilera paralela. Así el uso combinado de estas tecnologías permite una reducción sustancial de los plaguicidas y su impacto en el medio ambiente.

Agradecimientos

Los autores agradecen la financiación del presente trabajo a la COMISIÓN EUROPEA a través del proyecto RHEA. "Robot Fleets for Highly Effective Agriculture and Forestry Management" del 7 Programa Marco (proyecto nº 245986). Se agradece la contribución de todos los participantes en el proyecto: Agencia Estatal Consejo Superior de Investigaciones Científicas – CSIC-CAR (Centro de Automática y Robótica, Instituto de Ciencias Agrarias, Instituto de Agricultura Sostenible), CogVis GmbH, Forschungszentrum Telekommunikation Wien Ltd., Cyberbotics Ltd, Università di Pisa, Universidad Complutense de Madrid, Tropical, Soluciones Agrícolas de Precisión S.L., Universidad Politécnica de Madrid - UPM (ETS Ingenieros Agrónomos, ETS Ingenieros Industriales), AirRobot GmbH & Co. KG, Università degli Studi di Firenze, Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture -IRSTEA, CNH Belgium NV, CNH France SA, Bluebotics S.A. and CM Srl.

Bibliografía

- Brown, D. L., Giles, D. K., Oliver, M. N., and Klassen, P. (2008). Targeted spray technology to reduce pesticide in runoff from dormant orchards. *Crop Protection* **27**, 545-552.
- Correa, C., Valero, C., and Barriero, P. (2012). Characterization of vineyard's canopy through fuzzy clustering and svm over color images. In " IV International Workshop on Computer Image Analysis in Agriculture". Florentino Juste, Valencia.
- Chikamasa, T. (2012). Simulink Support for Kinect. pp. is a set of Simulink blocks which enables to make an interface between Simulink and a natural interaction device such as Microsoft Kinect. MathWorks, Natick, Massachusetts, Estados Unidos.
- Diago, M.-P., Correa, C., Millán, B., Barreiro, P., Valero, C., and Tardaguila, J. (2012). Grapevine Yield and Leaf Area Estimation Using Supervised Classification Methodology on RGB Images Taken under Field Conditions. *Sensors* **12**, 16988-17006.
- Gil, E., Escolà, A., Rosell, J. R., Planas, S., and Val, L. (2007). Variable rate application of plant protection products in vineyard using ultrasonic sensors. *Crop Protection* **26**, 1287-1297.
- Moltó, E., Martín, B., and Gutiérrez, A. (2001). PM—Power and Machinery: Pesticide Loss Reduction by Automatic Adaptation of Spraying on Globular Trees. *Journal of Agricultural Engineering Research* **78**, 35-41.
- Ongley, E. D. (1996). "Control of water pollution from agriculture," FAO, Rome.
- Pontikakos, C. M., Tsiligiridis, T. A., Yialouris, C. P., and Kontodimas, D. C. (2012). Pest management control of olive fruit fly (*Bactrocera oleae*) based on a location-aware agro-environmental system. *Computers and Electronics in Agriculture* **87**, 39-50.
- Reyes, J. F., Correa, C., Esquivel, W., and Ortega, R. (2012). Development and field testing of a data acquisition system to assess the quality of spraying in fruit orchards. *Computers and Electronics in Agriculture* **84**, 62-67.
- Siegfried, W., Viret, O., Huber, B., and Wohlhauser, R. (2007). Dosage of plant protection products adapted to leaf area index in viticulture. *Crop Protection* **26**, 73-82.
- Walklate, P. J., Cross, J. V., Richardson, G. M., Murray, R. A., and Baker, D. E. (2002). IT—Information Technology and the Human Interface: Comparison of Different Spray Volume Deposition Models Using LIDAR Measurements of Apple Orchards. *Biosystems Engineering* **82**, 253-267.